

SUBSTITUTE SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Au-Young, Janice
Lal, Preeti
Bandman, Olga
- (ii) TITLE OF THE INVENTION: TWO NEW HUMAN DNAJ-LIKE PROTEINS
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: To Be Assigned
 - (B) FILING DATE: Herewith
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/868,288
 - (B) FILING DATE: June 3, 1997
 - (A) APPLICATION NUMBER: 09/235,373
 - (B) FILING DATE: January 20, 1999
 - (A) APPLICATION NUMBER: 09/388,993
 - (B) FILING DATE: September 2, 1999
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Streeter, David G.
 - (B) REGISTRATION NUMBER: 43,168
 - (C) REFERENCE/DOCKET NUMBER: PF-0309-3 DIV
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 415-855-0555
 - (B) TELEFAX: 415-849-8886

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 358 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: SYNORAB01
 - (B) CLONE: 136466

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Ala Pro Gln Asn Leu Ser Thr Phe Cys Leu Leu Leu Tyr Leu
1 5 10 15

Ile Gly Ala Val Ile Ala Gly Arg Asp Phe Tyr Lys Ile Leu Gly Val
 20 25 30
 Pro Arg Ser Ala Ser Ile Lys Asp Ile Lys Lys Ala Tyr Arg Lys Leu
 35 40 45
 Ala Leu Gln Leu His Pro Asp Arg Asn Pro Asp Asp Pro Gln Ala Gln
 50 55 60
 Glu Lys Phe Gln Asp Leu Gly Ala Ala Tyr Glu Val Leu Ser Asp Ser
 65 70 75 80
 Glu Lys Arg Lys Gln Tyr Asp Thr Tyr Gly Glu Glu Gly Leu Lys Asp
 85 90 95
 Gly His Gln Ser Ser His Gly Asp Ile Phe Ser His Phe Phe Gly Asp
 100 105 110
 Phe Gly Phe Met Phe Gly Gly Thr Pro Arg Gln Gln Asp Arg Asn Ile
 115 120 125
 Pro Arg Gly Ser Asp Ile Ile Val Asp Leu Glu Val Thr Leu Glu Glu
 130 135 140
 Val Tyr Ala Gly Asn Phe Val Glu Val Val Arg Asn Lys Pro Val Ala
 145 150 155 160
 Arg Gln Ala Pro Gly Lys Arg Lys Cys Asn Cys Arg Gln Glu Met Arg
 165 170 175
 Thr Thr Gln Leu Gly Pro Gly Arg Phe Gln Met Thr Gln Glu Val Val
 180 185 190
 Cys Asp Glu Cys Pro Asn Val Lys Leu Val Asn Glu Glu Arg Thr Leu
 195 200 205
 Glu Val Glu Ile Glu Pro Gly Val Arg Asp Gly Met Glu Tyr Pro Phe
 210 215 220
 Ile Gly Glu Gly Glu Pro His Val Asp Gly Glu Pro Gly Asp Leu Arg
 225 230 235 240

 Phe Arg Ile Lys Val Val Lys His Pro Ile Phe Glu Arg Arg Gly Asp
 245 250 255
 Asp Leu Tyr Thr Asn Val Thr Val Ser Leu Val Glu Ser Leu Val Gly
 260 265 270
 Phe Glu Met Asp Ile Thr His Leu Asp Gly His Lys Val His Ile Ser
 275 280 285
 Arg Asp Lys Ile Thr Arg Pro Gly Ala Xaa Xaa Trp Lys Lys Gly Glu
 290 295 300
 Gly Leu Pro Asn Phe Asp Asn Asn Asn Ile Lys Gly Ser Leu Ile Ile
 305 310 315 320
 Thr Phe Asp Val Asp Phe Pro Lys Glu Gln Leu Thr Glu Glu Ala Arg
 325 330 335
 Glu Gly Ile Lys Gln Leu Leu Lys Gln Gly Ser Val Gln Lys Val Tyr
 340 345 350
 Asn Gly Leu Gln Gly Tyr
 355

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1376 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: SYNORAB01
- (B) CLONE: 136466

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCTCACCGGG	ACTCGGGACT	CCCGGGAAGT	GGACCGGCAG	AAGAGGGGGC	TAGCTAGCTG	60
TCTCTGCGGA	CCAGGGAGAC	CCCCGCGCCC	CCCCGGTGTG	AGGCGGCCTC	ACAGGGCCCG	120
GTGGGCTGGC	GAGCCGACGC	GGCGGCGGAG	GAGGCTGTGA	GGAGTGTGTG	GAACAGGACC	180
CGGGACAGAG	GAACCATGGC	TCCGCAGAAC	CTGAGCACCT	TTTGCTGTT	GCTGCTATAC	240
CTCATCGGGG	CGGTGATTGC	CGGACGAGAT	TTCTATAAGA	TCTTGGGGGT	GCCTCGAAGT	300
GCCTCTATAA	AGGATATTAA	AAAGGCCTAT	AGGAAACTAG	CCCTGCAGCT	TCATCCCGAC	360

```

CGGAACCCTG ATGATCCACA AGCCCAGGAG AAATTCCAGG ATCTGGGTGC TGCTTATGAG 420
GTTCTGTCAG ATAGTGAGAA ACGGAAACAG TACGATACTT ATGGTGAAGA AGGATTAAAA 480
GATGGTCATC AGAGCTCCCA TGGAGACATT TTTTCACACT TCTTTGGGGA TTTTGGTTTC 540
ATGTTTGGAG GAACCCCTCG TCAGCAAGAC AGAAATATTC CAAGAGGAAG TGATATTATT 600
GTAGATCTAG AAGTCACTTT GGAAGAAGTA TATGCAGGAA ATTTTGTGGA AGTAGTTAGA 660
AACAAACCTG TGGCAAGGCA GGCTCCTGGC AAACGGAAAGT GCAATTGTCTG GCAAGAGATG 720
CGGACCACCC AGCTGGGCCC TGGGCGCTTC CAAATGACCC AGGAGGTGGT CTGCGACGAA 780
TGCCCTAATG TCAAAC TAGT GAATGAAGAA CGAACGCTGG AAGTAGAAAT AGAGCCTGGG 840
GTGAGAGACG GCATGGAGTA CCCCTTTATT GGAGAAGGTG AGCCTCACGT GGATGGGGAG 900
CCTGGAGATT TACGGTTCCG AATCAAAGTT GTCAAGCACC CAATATTTGA AAGGAGAGGA 960
GATGATTTGT ACACAAATGT GACAGTCTCA TTAGTTGAGT CACTGGTTGG CTTTGAGATG 1020
GATATTACTC ACTTGGATGG TCACAAGGTA CATATTTCCC GGGATAAGAT CACCAGGCCA 1080
GGAGCGAANT ANTGGAAGAA AGGGGAAGGG CTCCCCAACT TTGACAACAA CAATATCAAG 1140
GGCTCTTTGA TAATCACTTT TGATGTGGAT TTTCCAAAAG AACAGTTAAC AGAGGAAGCG 1200
AGAGAAGGTA TCAAACAGCT ACTGAAACAA GGGTCAGTGC AGAAGGTATA CAATGGACTG 1260
CAAGGATATT GAGAGTGAAT AAAATTGGAC TTTGTTTAAA ATAAGTGAAT AAGCGATATT 1320
TATTATCTGC AAGGTTTTTT TGTGTGTGTT TTTGTTTAA TTTTCAATAT GCAAGT 1376

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 330 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: HNT2RAT01
- (B) CLONE: 260873

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Val Asp Tyr Tyr Glu Val Leu Gly Val Gln Arg His Ala Ser Pro
 1          5          10          15
Glu Asp Ile Lys Lys Ala Tyr Arg Lys Leu Ala Leu Lys Trp His Pro
 20          25          30
Asp Lys Asn Pro Glu Asn Lys Glu Glu Ala Glu Arg Lys Phe Lys Gln
 35          40          45
Val Ala Glu Ala Tyr Glu Val Leu Ser Asp Ala Lys Lys Arg Asp Ile
 50          55          60
Tyr Asp Lys Tyr Gly Lys Glu Gly Leu Asn Gly Gly Gly Gly Gly Gly
 65          70          75          80
Ser His Phe Asp Ser Pro Phe Glu Phe Gly Phe Thr Phe Arg Asn Pro
 85          90          95
Asp Asp Val Phe Arg Glu Phe Phe Gly Gly Arg Asp Pro Phe Ser Phe
100          105          110
Asp Phe Phe Glu Asp Pro Phe Glu Asp Phe Phe Gly Asn Arg Arg Gly
115          120          125
Pro Arg Gly Ser Arg Ser Arg Gly Thr Gly Ser Phe Phe Ser Ala Phe
130          135          140
Ser Gly Phe Pro Ser Phe Gly Ser Gly Phe Ser Ser Phe Asp Thr Gly
145          150          155          160
Phe Thr Ser Phe Gly Ser Leu Gly His Gly Gly Leu Thr Ser Phe Ser
165          170          175
Ser Thr Ser Phe Gly Gly Ser Gly Met Gly Asn Phe Lys Ser Ile Ser
180          185          190
Thr Ser Thr Lys Met Val Asn Gly Arg Lys Ile Thr Thr Lys Arg Ile
195          200          205
Val Glu Asn Gly Gln Glu Arg Val Glu Val Glu Glu Asp Gly Gln Leu
210          215          220
Lys Ser Leu Thr Ile Asn Gly Val Ala Asp Asp Ala Leu Xaa Glu
225          230          235          240
Glu Arg Met Arg Arg Gly Gln Asn Val Leu Pro Ala Gln Pro Ala Gly
245          250          255
Leu Arg Pro Pro Lys Pro Pro Arg Pro Ala Ser Leu Leu Arg His Xaa
260          265          270

```

Pro His Cys Leu Ser Lys Glu Glu Gly Glu Gln Asp Arg Pro Trp Ala
 275 280 285
 Pro Xaa Xaa Trp Xaa Pro Leu Ala Ser Xaa Ala Gly Xaa Xaa Glu Gly
 290 295 300
 Xaa Lys Arg Met Xaa Ala Glu Ala Glu Arg Gly Val Glu Glu Glu Glu
 305 310 315 320
 Val Asp Gln Arg Gln Ser Leu Asp Arg Thr
 325 330

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1330 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: HNT2RAT01
- (B) CLONE: 260873

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGNAGGAGAG	NAAAGGAAAG	NCGCCGCAGG	AGCCGCCGCN	ACCACCAGCG	NCACANTCCT	60
GGNGCTNTGA	GGAGATTCGG	GCCGTCACCC	TGCCTCCCCT	GCTTCCCGCC	ACCGGCCGCT	120
TCTTTCCTCG	GACCCATTCC	AACAATCTCG	TAAAAATG	TGGATTACTA	TGAAGTTCTA	180
GGCGTGCAGA	GACATGCCTC	ACCCGAGGAT	ATTAAAAAG	CATATCGGAA	ACTGGCACTG	240
AAGTGGCATC	CAGATAAAAA	TCCTGAGAAT	AAAGAAGAAG	CAGAGAGAAA	ATTCAAGCAA	300
GTAGCGGAGG	CATATGAAGT	GCTGTCCGAT	GCTAAGAAAC	GGGACATCTA	TGACAAATAT	360
GGCAAAGAAG	GATTAAATGG	TGGNGGNGGN	GGTGGAAAGT	ATTTTGACAG	TCCATTTGAA	420
TTTGGCTTCA	CATTCCGTAA	CCCAGATGAT	GTCTTCAGGG	AATTTTTTGG	TGGAAGGGAC	480
CCATTTTCAT	TTGACTTCTT	TGAAGACCTT	TTTGAGGACT	TCTTTGGGAA	TGGAAGGGGT	540
CCCCGAGGAA	GCAGAAGCCG	AGGGACGGGG	TCGTTTTTCT	CTGCGTTCAG	TGGATTTCCG	600
TCTTTTGAA	GTGGATTTTC	TTCTTTTGAT	ACAGGATTTA	CTTCATTTGG	GTCAC TAGGT	660
CACGGGGGCC	TCACCTTCAT	CTCTTCCACG	TCATTTGGTG	GTAGTGGCAT	GGGCAACTTC	720
AAATCGATAT	CAACTTCAAC	TAAAATGGTT	AATGGCAGAA	AAATCACTAC	AAAGAGAATT	780
GTCGAGAACG	GTCAAGAAAG	AGTAGAAGTT	GAAGAAGATG	GCCAGTTAAA	GTCCTTAACA	840
ATAAATGGTG	TKGCCGACGA	CGATGCCCTC	GSTGAGGAGC	GCATGCGGAG	AGGCCAGAAC	900
GTCTTGCCAG	CCCAGCCTGC	CGGCCTCCGA	CCGCCGAAGC	CGCCCCGGCC	TGCCTCGTTG	960
CTGAGACACG	NGCCTCATTG	TCTCTCTAAG	GAGGAGGGCG	AGCAGGACCG	ACCTTGGGCA	1020
CCCGNGNCCCT	GGNNCCCCCT	CGCTTCCNCA	GCAGGNTTNN	AAGAAGGTNG	CAAGAGGATG	1080
NAAGCAGAAG	CAGAGAGAGG	AGTCGAAGAA	GAAGAAGTCG	ACCAAAGGCA	ATCACTAGAC	1140
CGGACTTGAG	GCACGCGGTG	CACCCCCAGA	CGCTGGCGCT	CCACCGTGCT	CGGCATGCGG	1200
TCGTGCACAC	GCGCTAGGTA	GCAGCGTCGG	TCAGGACTGT	CTCGAGGCCA	CACTCGCTCG	1260
GCAGGATTAT	GCGATCACGG	ATCAGTCAGA	GCAGGGTCAG	GAGACGGGGC	TGACGGCACG	1320
GGTGGCGGGG						1330

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 397 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 306714

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Val Lys Glu Thr Thr Tyr Tyr Asp Val Leu Gly Val Lys Pro Asn
 1 5 10 15
 Ala Thr Gln Glu Glu Leu Lys Lys Ala Tyr Arg Lys Leu Ala Leu Lys

(2) INFORMATION FOR SEQ ID NO:6:

(A) LENGTH: 351 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) LIBRARY: GenBank
(B) CLONE: 32469

Met Ala Ser Tyr Tyr Glu Ile Leu Asp Val Pro Arg Ser Ala Ser Ala
1 5 10 15
Asp Asp Ile Lys Lys Ala Tyr Arg Arg Lys Ala Leu Gln Trp His Pro

(2) INFORMATION FOR SEQ ID NO:7:

(A) LENGTH: 277 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) LIBRARY: GenBank
(B) CLONE: 32470

Met	Ala	Ser	Tyr	Tyr	Glu	Ile	Leu	Asp	Val	Pro	Arg	Ser	Ala	Ser	Ala
1				5					10					15	
Asp	Asp	Ile	Lys	Lys	Ala	Tyr	Arg	Arg	Lys	Ala	Leu	Gln	Trp	His	Pro
			20					25					30		
Asp	Lys	Asn	Pro	Asp	Asn	Lys	Glu	Phe	Ala	Glu	Lys	Lys	Phe	Lys	Glu
		35				40						45			
Val	Ala	Glu	Ala	Tyr	Glu	Val	Leu	Ser	Asp	Lys	His	Lys	Arg	Glu	Ile
	50					55					60				
Tyr	Asp	Arg	Tyr	Gly	Arg	Glu	Gly	Leu	Thr	Gly	Thr	Gly	Thr	Gly	Pro
65					70					75					80

Ser	Arg	Ala	Glu	Ala	Gly	Ser	Gly	Gly	Pro	Gly	Phe	Thr	Phe	Thr	Phe
				85					90					95	
Arg	Ser	Pro	Glu	Glu	Val	Phe	Arg	Glu	Phe	Phe	Gly	Ser	Gly	Asp	Pro
			100					105					110		
Phe	Ala	Glu	Leu	Phe	Asp	Asp	Leu	Gly	Pro	Phe	Ser	Glu	Leu	Gln	Asn
		115					120					125			
Arg	Gly	Ser	Arg	His	Ser	Gly	Pro	Phe	Phe	Thr	Phe	Ser	Ser	Ser	Phe
	130					135					140				
Pro	Gly	His	Ser	Asp	Phe	Ser	Ser	Ser	Ser	Phe	Ser	Phe	Ser	Pro	Gly
145					150					155					160
Ala	Gly	Ala	Phe	Arg	Ser	Val	Ser	Thr	Ser	Thr	Thr	Phe	Val	Gln	Gly
			165					170						175	
Arg	Arg	Ile	Thr	Thr	Arg	Arg	Ile	Met	Glu	Asn	Gly	Gln	Glu	Arg	Val
			180				185						190		
Glu	Val	Glu	Glu	Asp	Gly	Gln	Leu	Lys	Ser	Val	Thr	Ile	Asn	Gly	Val
	195						200					205			
Pro	Asp	Asp	Leu	Ala	Arg	Gly	Leu	Glu	Leu	Ser	Arg	Arg	Glu	Gln	Gln
	210					215					220				
Pro	Ser	Val	Thr	Ser	Arg	Ser	Gly	Gly	Thr	Gln	Val	Gln	Gln	Thr	Pro
225					230					235					240
Ala	Ser	Cys	Pro	Leu	Asp	Ser	Asp	Leu	Ser	Glu	Asp	Glu	Asp	Leu	Gln
			245						250					255	
Leu	Ala	Met	Ala	Tyr	Ser	Leu	Ser	Glu	Met	Glu	Ala	Ala	Gly	Lys	Lys
			260					265					270		
Pro	Ala	Asp	Val	Phe											
		275													